

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра молекулярной биологии

Аннотация к магистерской диссертации

Масный

Александр Александрович

АННОТАЦИЯ ГЕНОМА *PESTOBACTERIUM CAROTOVORUM* 3-2

Научный руководитель:

доцент кафедры молекулярной биологии

Евгений Артурович Николайчик, Ph. D.

Минск, 2015

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

АННОТАЦИЯ БАКТЕРИАЛЬНОГО ГЕНОМА, "РУЧНАЯ" АННОТАЦИЯ, АВТОМАТИЧЕСКАЯ АННОТАЦИЯ, *PECTOBACTERIUM CAROTOVORUM* 3-2, *PECTOBACTERIUM ATROSEPTICA* SCRI1043, *DICKEYA DADANTII* 3937, *ESCHERICHIA COLI* K-12 MG1655.

Объект исследования – секвенированный геном *Pectobacterium carotovorum* штамма 3-2, который был аннотирован программными средствами.

Предмет исследования – аннотация бактериального генома, содержащая ошибки и неточности, которые требуется исправить при ручном аннотировании генома.

Целью данной работы стало повышение качества автоматической аннотации программными средствами с использованием молекулярных баз данных.

Использование информации из экспериментальных статей позволило внести коррекцию в 9 локусов пектолитических ферментов, в 7 локусов, связанных утилизацией галактана и 8 – с утилизацией β -глюкозидов. Верификация на основании публикаций по сравнительной геномике *Pectobacterium atroseptica* SCRI1043 и *Dickeya dadantii* 3937 позволила исправить аннотацию в 4 локусах консервативных генов и в 14 локусах, связанных с утилизацией железа.

Сравнение с референсными геномами: *Pectobacterium atroseptica* SCRI1043, *Dickeya dadantii* 3937, *Escherichia coli* K-12 субштамм MG1655 позволило дополнительно исправить аннотацию около 900 локусов, а именно, были исправлены названия 763 генов и 373 продуктов генов, добавлены ЕС-номера у 113 ферментов, исправлены 70 заметок о продукте гена и удалены 69 некорректных структурных элемента генома.

Магистерская диссертация представлена в 3 главах. В первой главе, обзор литературы, описан процесс аннотации бактериального генома, а также описаны свойства организмов, геномы которых использовались в работе. Во второй главе идёт речь о материалах и методах исследования. В последней главе обсуждаются полученные результаты.

Магистерская диссертация: 59 страниц, 12 рисунков, 11 таблиц, 66 источников.

АГУЛЬНАЯ ХАРАКТАРЫСТЫКА ПРАЦЫ

АНАТАЦЫЯ БАКТЭРЫЯЛЬНАГА ГЕНОМУ, "РУЧНАЯ" АНАТАЦЫЯ, АЎТАМАТЫЧНАЯ АНАТАЦЫЯ, *PECTOBACTERIUM CAROTOVORUM* 3-2, *PECTOBACTERIUM ATROSEPTICA* SCRI1043, *DICKEYA DADANTII* 3937, *ESCHERICHIA COLI* K-12 MG1655.

Аб'ект даследавання – секвенаваны геном *Pectobacterium carotovorum* штама 3-2, які быў анатаваны праграмнымі сродкамі.

Прадмет даследавання – анатацыя бактэрыяльнага геному, што змяшчае памылкі і недакладнасці, якія патрабуюцца выправіць пры ручным анатаванні геному.

Мэтай дадзенай працы стала падвышэнне якасці аўтаматычнай анатацыі праграмнымі сродкамі з выкарыстаннем малекулярных баз даных.

Выкарыстанне інфармацыі з эксперыментальных артыкулаў дазволіла ўнесці карэкцыю ў 9 локусаў пекталітычных ферментаў, у 7 локусаў, злучаных з утылізацыяй галактану і 8 – з утылізацыяй β-глюказідаў. Верыфікацыя на падставе публікацый па параўнальнай геноміцы *Pectobacterium atroseptica* SCRI1043 і *Dickeya dadantii* 3937 дазволіла выправіць анатацыю ў 4 локусах кансерватыўных генаў і ў 14 локусах, злучаных з утылізацыяй жалеза.

Параўнанне з рэферэнсавымі геномамі: *Pectobacterium atroseptica* SCRI1043, *Dickeya dadantii* 3937, *Escherichia coli* K-12 суб штама MG1655, дазволіла дадаткова выправіць анатацыю каля 900 локусаў, менавіта: былі выпраўлены назвы 763 генаў і 373 прадуктаў генаў, дададзены ЕС-нумары ў 113 ферментаў, выпраўлены 70 нататкаў пра прадукт гена і выдалены 69 некарэктных структурных элемента геному.

Магістарская дысертацыя пададзена ў 3 частках. У першай частцы, агляд літаратуры, апісаны працэс анатацыі бактэрыяльнага геному, а таксама апісаны ўласцівасці арганізмаў, геномы якіх выкарыстоўваліся ў працы. У другой частцы ідзе гаворка пра матэрыялы і метады даследавання. У апошняй частцы абмяркоўваюцца атрыманыя вынікі.

Магістарская дысертацыя: 59 старонак, 12 малюнкаў, 11 табліц, 66 крыніц.

ZUSAMMENFASSUNG

ANNOTATION DES BAKTERIELLEN GENOMES, "MANUELLE" ANNOTATION, AUTOMATISCHE ANNOTATION, *PECTOBACTERIUM CAROTOVORUM* 3-2, *PECTOBACTERIUM ATROSEPTICA* SCRI1043, *DICKEYA DADANTII* 3937, *ESCHERICHIA COLI* K-12 MG1655.

Das Objekt der Forschung ist das sequenzierte Genom von *Pectobacterium carotovorum* Stamm 3-2, das durch die Software annotiert war.

Der Gegenstand der Forschung ist die bakterielle Genom-Annotation, die Fehler und Ungenauigkeiten enthält und muss durch die manuelle Annotation des Genoms repariert werden.

Das Ziel dieser Arbeit ist die Qualität der automatischen Annotation mithilfe von Software und von der molekularen Datenbasen zu verbessern.

Unter Zuhilfenahme von experimentellen Artikeln gestattet die Korrektur der

9 Loci der pektolytische Enzyme, der 7 Loci, die mit Nutzung der Galactane und der 8 Loci, die mit Nutzung der β -Glycoside assoziiert sind. Überprüfung auf der Grundlage der Publikationen des vergleichenden Genomik von *Pectobacterium atroseptica* SCRI1043 und *Dickeya dadantii* 3937 ermöglicht die Verbesserung der 4 konservativen Genen und der 14 Loci, die mit der Verwertung von Eisen zugeordnet, zu korrigieren.

Vergleich mit der Referenzgenomen, nämlich *Pectobacterium atroseptica* SCRI1043, *Dickeya dadantii* 3937 und *Escherichia coli* K-12 MG1655, erlaubt zusätzlich die korrekte Annotation der circa 900 Loci zu machen, und zwar 763 Namen der Genen, 373 Genprodukte, die EG-Nummern von 113 Enzymen, 70 Notizen von dem Produkt des Gens sind korrigiert und 69 fehlerhaft Features des Genoms sind gelöscht.

Masterarbeit ist in drei Kapiteln vorgestellt. Im ersten Kapitel werden die theoretischen Überblick über die Annotation des Genoms sowie Organismen, deren Genome in die Arbeit verwendet sind, diskutiert. Das zweite Kapitel enthält die Information über die Materials und Methoden der Forshung. Das letzte Kapitel werden die Untersuchungsergebnisse beschrieben.

Masterarbeit: 59 Seiten, 12 Abbildungen, 11 Tabellen, 66 Quellen.